DATA LEAKAGE CULPRIT DETECTION USING ENCRYPTION TECHNIQUES

*Akshay Kulkarni [1] | Ashwini Wangate [1] | Nischal Kamble [1] | Priyanka Chinchole [1]

[1] Student, Information Technology, SKNCOE, Pune, India - 411041.*Corresponding Author

ABSTRACT

Data leakage is a big challenge as critical organizational data should be protected from unauthorized access. Data leakage is defined as the distribution of private or sensitive data to unauthorized entity accidentally or unintentionally [1]. Sensitive data of companies and organizations includes intellectual property (IP), financial information, patient information, personal credit-card data, and other information depending on the business and the industry. This increases the risk of confidential information falling into unauthorized hands. Whether caused by malicious intent, or an inadvertent mistake, by an insider or outsider, exposed sensitive information can seriously hurt an organization. It is very hard for any system administrator to trace out the data leaker among the system users. It creates a lot many ethical issues in the working environment of the office. In the recent years internet technologies has become the backbone of any business organization. These organizations use this facility to improve their efficiency by transferring data from one location to another. But, there are number of threats in transferring critical organizational data as any culprit employee may public this data. This problem is known as data leakage problem. In the proposed work, we are suggesting a model for data leakage problem. In this model, our aim is to identify the culprit who has leaked the critical organizational data.

KEYWORDS: Data leakage.

1. Introduction:

Every organization shares data digitally among its employees and customers. Furthermore, in many cases, sensitive data is shared among various stakeholders such as employees working from outside the organizational premises (e.g., on laptops), business partners and customers. Data leakage is a big challenge in front of the industries & different institutes. Though there are number of systems designed for the data security by using different encryption algorithms, there is a big issue of the integrity of the users of those systems. It is very hard for any system administrator to trace out the data leaker among the system users. It creates a lot many ethical issues in the working environment of the office. The data leakage detection industry is very heterogeneous as it evolved out of ripe product lines of leading IT security vendors. A broad arsenal of enabling technologies such as firewalls, encryption, access control, identity management, machine learning content/context based detectors and others have already been incorporated to offer protection against various facets of the data leakage threat.

Consider an organization where sensitive data is shared amongst its employees, if any employee leaks this sensitive information to rival organizations it can be a big problem. It is very difficult to identify the culprit who leaked the data as there are many suspects. Our proposed idea gives a solution to this problem. The main question at hand is how to differentiate the culprit from other employees.

Traditionally, leakage detection is handled by watermarking, e.g., a unique code is embedded in each distributed copy. If that copy is later discovered in the hands of an unauthorized party, the leaker can be identified [3]. Watermarks can be very useful in some cases, but again, involve some modification of the original data. Furthermore, watermarks can sometimes be destroyed if the data recipient is malicious.One solution to this can be achieved using encryption techniques. We use encryption techniques to embed a key in the video and keep a record of it. Whenever a leaked video is encountered, we can match the key in the video and the database to find the respective user who is the culprit.

2. Materials and Methods:

For the sake of simplicity let us say that an organization distributes videos to its employees. This video data is sensitive and should not fall in wrong hands. A video can be downloaded by many users. If any user leaks this information, it is possible to identify him/her. We propose to embed an unique key in the video every time a user downloads the data, a database should be maintained containing the list of assigned keys to the corresponding videos and the users that downloaded it. This database stores the user and the key assigned to that particular user while downloading the data. Now we have a database that provides us with the user who downloaded the information. Every time a video is downloaded, a new key is generated so even if the same user downloads the same video, a new entry will be made in the database.

Assume that we have acquired the leaked data, now we can retrieve the key in the video using decryption and compare it in our database. The key that matches provides us with the user corresponding to the key, which is our culprit. Further legal action can be taken on the user.

When considering video data, we can use steganography techniques to embed the key in the video. The encryption and decryption of the key generated can be handled using RSA algorithm.

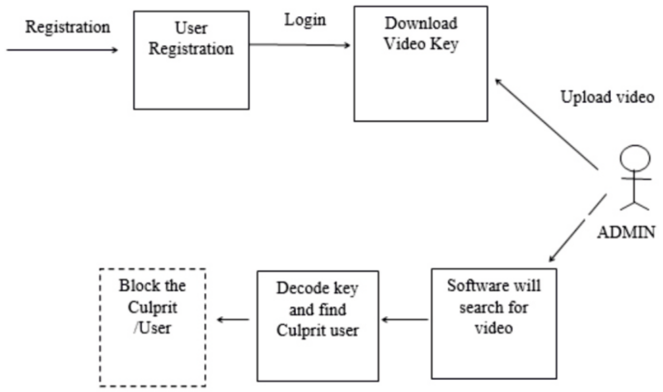

Fig 1. System design

3. Results:

Our implementation shows that this technique can be used to successfully encrypt a key and cross check the key in the database to identify the data leakage culprit. While considering video data care should be taken to use lossless compression techniques, lossy compression techniques may remove video frames that contain embedded key.

Tables and figures:

SR.No	Name	Page No
1	System design	2

REFERENCES:

1. Neeraj Kumar, Vijay Katta, Himanshu Mishra, Hitendra Garg. (2014). Detection of Data Leakage in Cloud Computing Environment. IEEE. Sixth International Conference on Computational Intelligence and Communication Networks.
2. Panagiotis Papadimitriou, Hector Garcia-Molina. (2009). A Model for Data Leakage Detection, IEEE.International Conference on Data Engineering.
3. Panagiotis Papadimitriou, and Hector Garcia-Molina.(2011).Data Leakage Detection, IEEE transactions on knowledge and data engineering.
4. S.Umamaheswari, H.Arthi Geetha. (2011). Detection of Guilty Agents.IEEE. Proceedings of the National Conference on Innovations in Emerging Technology.